MPM1D **4.1 Scatter Plots** **Date:** _____

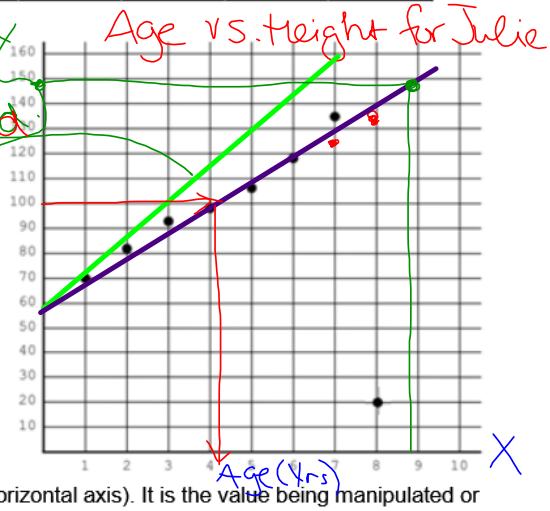
A scatter plot is graph that shows the relationship between 2 variables. The points on the scatter plot show a pattern or a trend. From the pattern or trend you can describe the correlation.

Example: Julie gathered information about her age and height from the markings on the wall in her house and produced the scatter plot below. Complete the graph by adding titles and labeling axis.

Age (years)	1	2	3	4	5	6	7	8
Height (cm)	70	82	93	98	106	118	127	135

Good Graphing Standards Check list

1. Title "X" vs "Y"
2. Axis labelled with units in brackets
3. 2/3 graph used
4. Line/curve dotted outside collected data



Questions

a. Describe the trend in the data.

Julie is getting taller

b. Describe the relationship.

As Julie's age increases, her height increases.

X-axis Y axis

Variables and Axis

- The *independent variable* is located on the X axis (horizontal axis). It is the value being manipulated or changed by the experimenter. In this scatter plot the Age is the independent variable.
- The *dependent variable* is located on the Y axis (vertical axis). It is the observed result of the independent variable. In this scatter plot the height is the dependent variable.

Line of Best Fit

To be able to make predictions, we need to model the data with a line or a curve of Best fit. To draw a line of best fit:

Note:
The independent variable comes *first* in the table of values.

1. The line should go through as many points as possible. (doesn't have to go any)
2. There should be the same number of point above and below the line.

Outliers

An outlier is a piece of data that does not seem to follow the trend demonstrated by the rest of the data set. It is important to identify outliers before drawing a line of best fit as they should be excluded from the model. The coordinates of the outlier in this example are (8, 20).

Interpolating vs. Extrapolating

When you are interpolating you are making predictions within the data set. When you are extrapolating you are making predictions outside the data set. Predictions made using interpolation are more accurate.

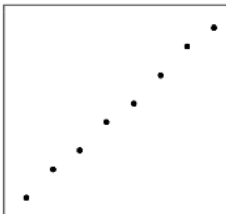
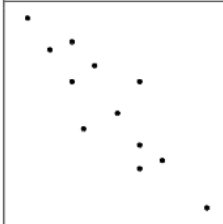
Hint:
You are interpolating when the value you are finding is somewhere between the first point and the last point.

Making Predictions

You can use a line of best fit to make predictions.

Question	Answer	Method of Prediction (Interpolation/Extrapolation)
How tall was Julie when she was 5 years old?	106cm	Interpolation
How tall will Julie be when she is 9 years old?	150cm	Extrapolate
How old was Julie at 100 cm tall?	4yrs	Interpolation
How tall was Julie when she was born?	55cm	Extrapolation

Correlation

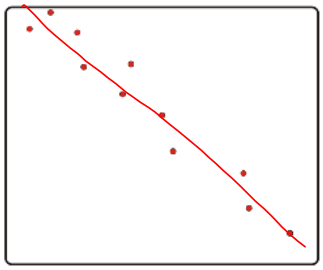
	A scatter plot shows a POSITIVE correlation when the pattern rises up to the right. This is sometimes called a ~~strong~~ relationship. (dots are really close together) *This means that the two quantities increase together.*
	A scatter plot shows a Negative correlation when the pattern falls down to the right. This is sometimes called an weak relationship. *This means that as one quantity increases the other decreases.*
	A scatter plot shows NO correlation when no pattern appears. *Hint:* *If the points are roughly enclosed by a circle, then there is no correlation.*

If the points nearly form a line, then the correlation is STRONG. If the points are dispersed more widely, but still form a rough line, then the correlation is WEAK.

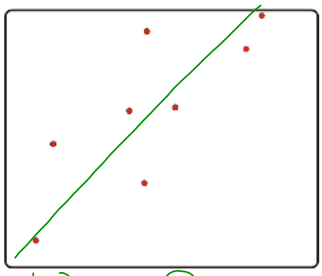
Hint:
To visualize this, enclose the plotted points in an oval.
If the oval is *thin*, then the correlation is *strong*.
If the oval is *fat*, then the correlation is *weak*.

Draw a line/curve of best fit for each of the scatter plots below. Write two or three key words to describe each relation on the line below the scatter plot.

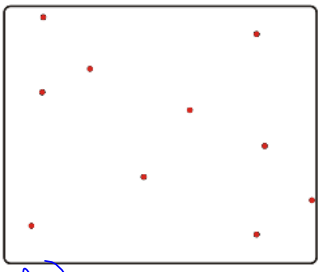
- *rises upward to the right*
- *falls downward to the right*
- *no relationship*
- *strong, weak*
- *linear, non-linear*



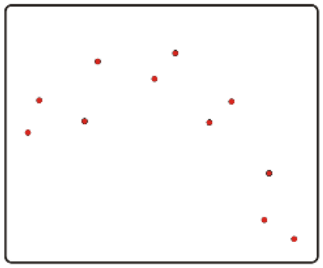
a) Strong, negative



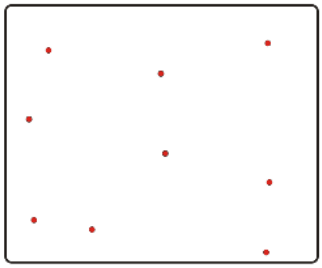
b) Weak Positive



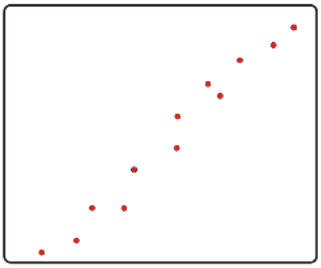
c) No correlation



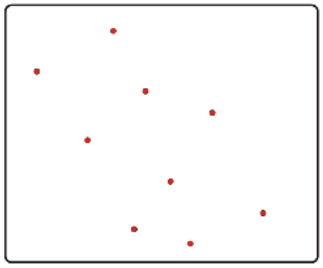
d) _____



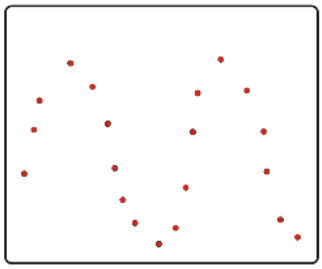
e) No correlation



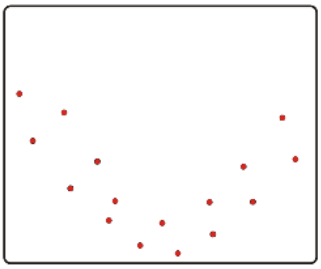
f) _____



g) _____



h) _____



i) _____

EXAMPLE 1

Question: Are a person's age and earnings related?

Hypothesis:

Determining your Scale

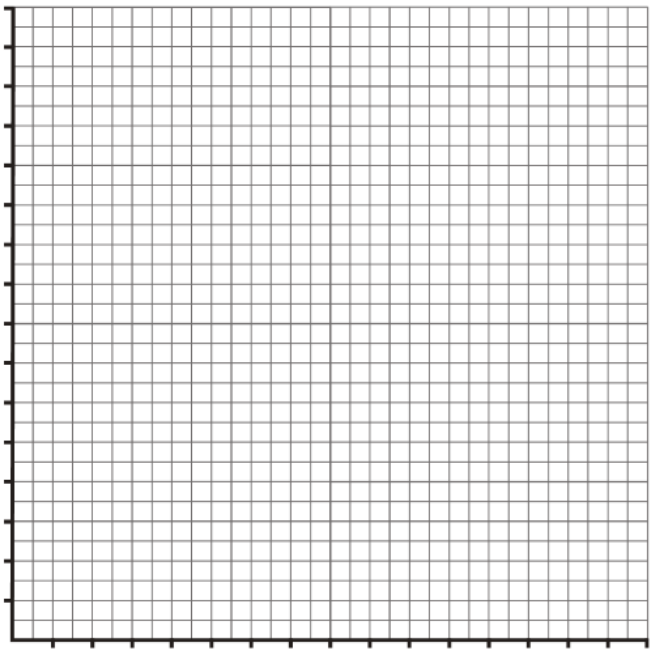
The symbol _____ is used to signal a "break" in the axis when the scale does not start at zero to avoid large empty space in one corner of the graph. When determining your scale follow these steps:

1. Count the number of squares on your axis.
2. Find the difference between your highest and lowest value.
3. Divide Step 2 by Step 1
4. Choose a whole number larger than this value.
5. Data must take up 2/3 of the graph.

x-axis	y-axis

Plot the data on the scatter plot below and draw a line of best fit.

Age	Earnings ($)
25	22000
30	26500
35	29500
37	29000
38	30000
40	32000
41	35000
45	36000
55	41000
60	41000
62	42500
65	43000



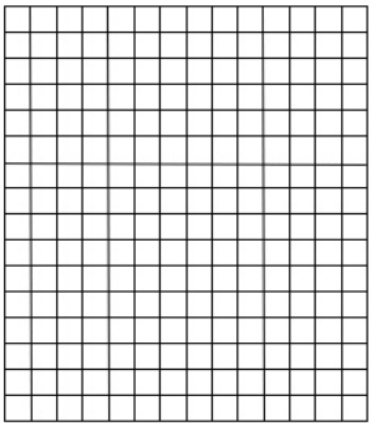
Questions – To be done on a separate sheet

1. Which is the dependent and which is the dependent variable? How do you know?
2. Describe the relationship between the 2 variables.
3. Would this relationship be true of all values? Explain.
4. Explain what the point (41, 35000) represents.
5. Use your graph to determine how much this person made at 50. Show this on your graph.
6. When did this person make $20 000. Show this on your graph.

EXAMPLE 2

Anthropologists and forensic scientists use data to determine information about people. Scientists can make predictions about the height, age, and sex of the person they are examining by looking for relationships in large amounts of data.

1. Construct a graph of the length of the humerus bone vs. the length of the radius.



Length of Radius (cm)	Length of Humerus (cm)
25	29.7
22	26.5
23.5	27.1
22.5	26
23	28
22.6	25.2
21.4	24
21.9	23.8
23.5	26.7
24.3	29
24	27

2. Circle the point on the graph that represents the data for a radius that is 21.9 cm long.
   How long is the humerus? _____.

3. Put a box around the point on the graph that represents the data for a humerus that is 27.1 cm long. How long is the radius? _____.

4. Describe the trend.

5. Describe the relationship: As the length of the radius gets longer, the humerus _____.

6. a) Draw a line of best fit.

   b) Use the line of best fit to predict the length of the humerus, if the radius is 24.5 cm long. Show this on your graph. Did you interpolate or extrapolate?

   c) Use the line of best fit to predict the length of the radius, if the humerus is 25 cm long. Show this on your graph. Did you interpolate or extrapolate?